Model Evaluation Status

Meeting of GE and EPA

Montvale, NJ

September 27, 2010

LimnoTech
Water | Environment | Scientists | Engineers

Overview

- Model evaluation objectives
- Plan to meet objectives
- Activities underway
- Issues identified to date
  - Illustrative examples, all draft/work in progress
  - Promote dialogue
- Path to peer review and acceptance
- Opportunities for and value of collaboration

Management-Driven Objectives

- Level 1: Comparing system recovery from dredging impacts to MNA without dredging
- Level 2: Aiding in the design and optimization of the Phase 2 dredging program
- Level 3: Using models to make adjustments in the field during Phase 2, in response to monitoring data

Fundamental Evaluation Criterion

- Is the suite of models consistent with and sufficiently constrained by site data to support management decisions?
  - Baseline model (historical/MNA) - does it accurately reproduce:
    - High-flow event outcomes
    - Low-flow conditions
    - Long-term trends in exposures
  - Dredging model – can the fate of resuspended PCBs be reliably predicted, based on Phase 1 model-data comparisons?

Activities

- Transfer of models from AQ
- Evaluation of data used
- Review of model inputs
- Benchmarking of model outputs
- Verification of model calculations
- Model-data comparisons
- Diagnostic evaluations

Siginificant Issues Identified

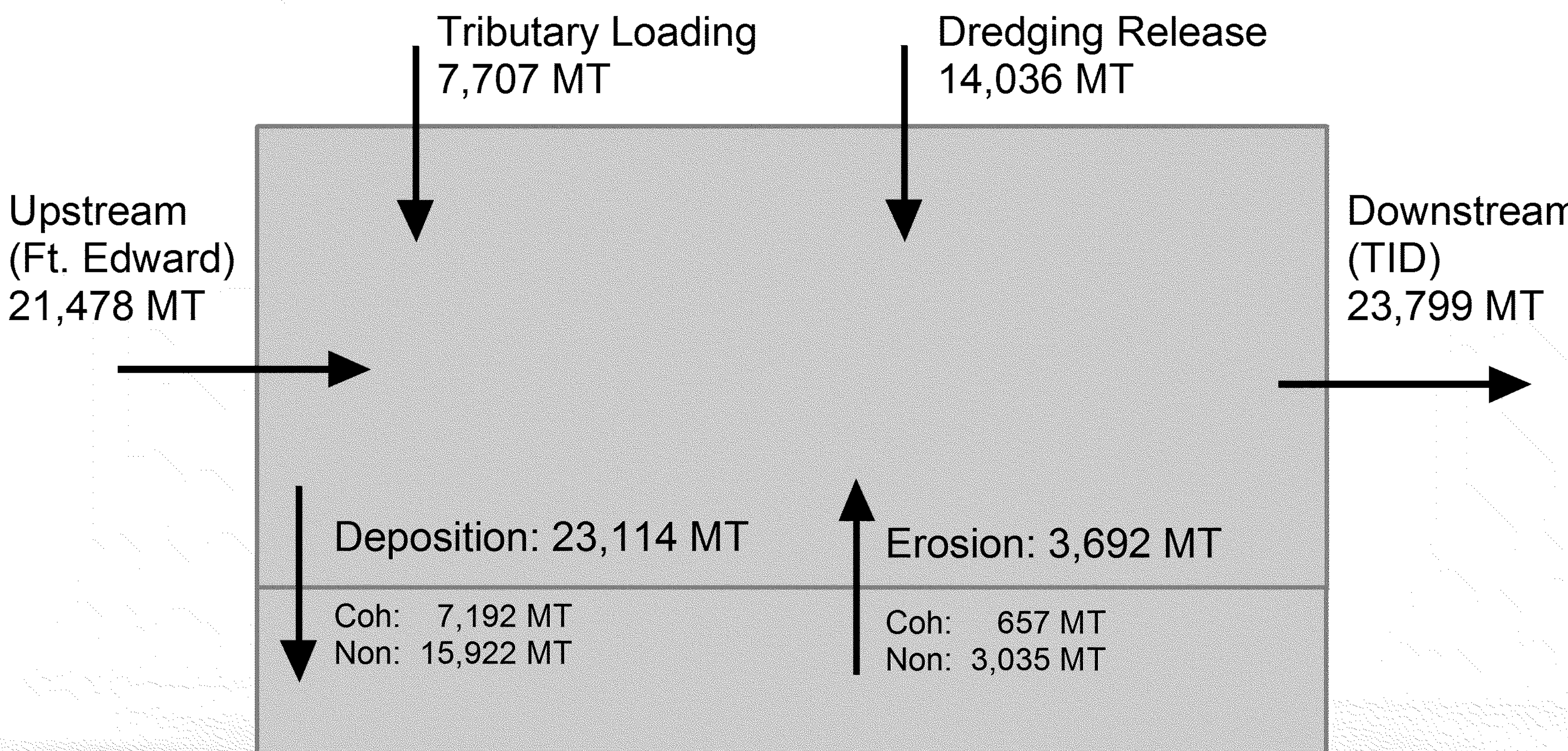
- Dredging simulations
  - Treatment of resistance to desorption
  - Predicted fate of resuspended PCBs is uncertain & inconsistent with Phase 1 near-field data
- Baseline (historical & MNA) simulations
  - Rigorous model-data comparisons needed
  - Predicted long-term MNA trends not sufficiently tested against data
- Sediment transport model assumptions deviate from site data

Challenge – Modeling Dredged Solids vs. Other Solids

- Three types:
  - Watershed solids, present at low flow
  - Flow-resuspended solids
  - Dredge-resuspended solids
- All are present (e.g. during 2009 simulation) and they may transport PCBs differently
  - Watershed solids have had time adsorb/desorb
  - Resuspended solids may be slow to desorb, and the two types may do so at different rates

Solids Mass Balance: Reach 8, 2009

(5% dredge release scenario)



How Resistant Sorption Is Modeled

- Baseline simulations
  - PCB erosion flux is reduced by half for coarse particles
  - No special treatment for suspended particles
- Dredging simulations
  - Two model runs are averaged, one run assuming all water column PCBs sorb tightly to solids, the other run assuming normal partitioning
  - Applied to all PCBs and solids in the river

Comparison – MNA 2009

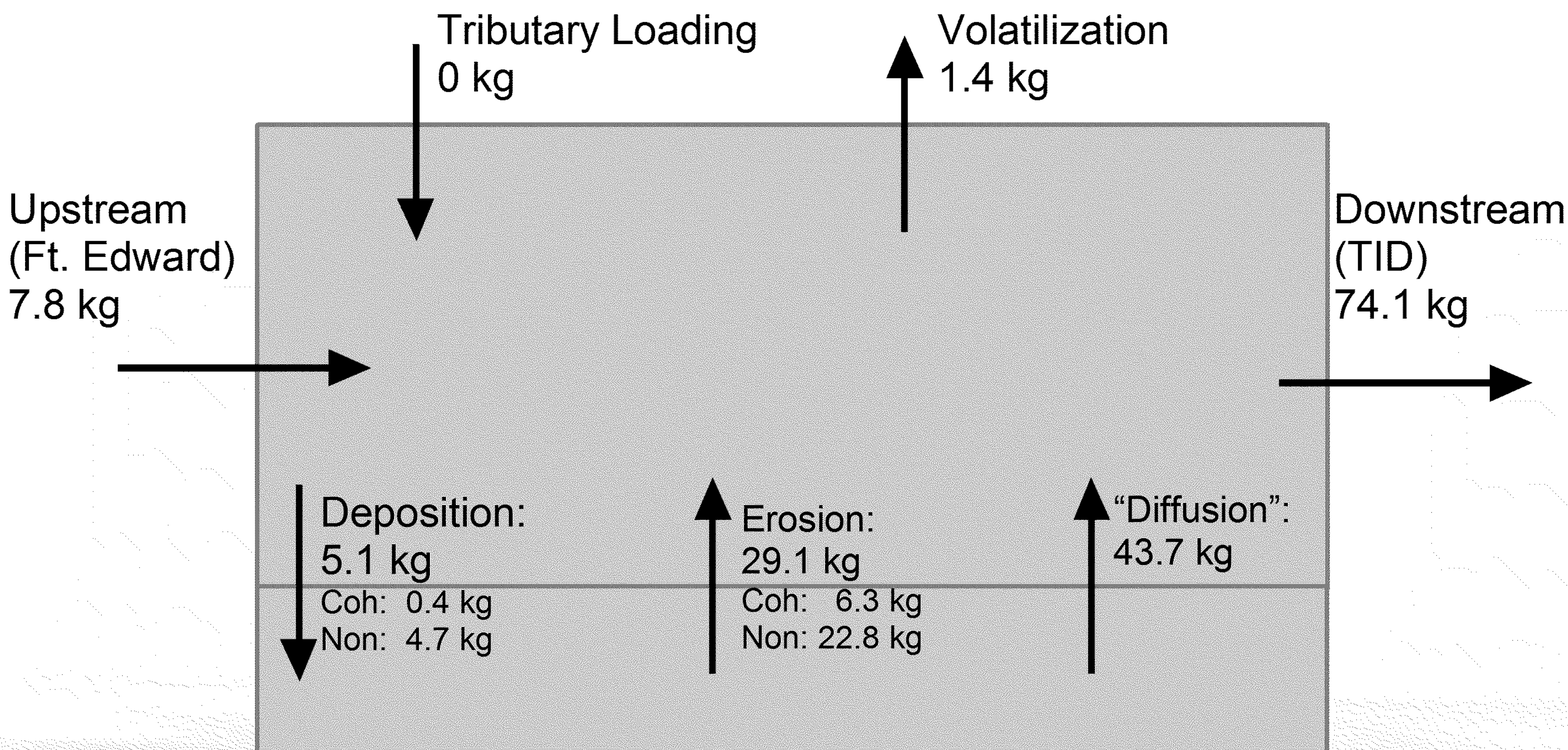
- What's the effect of the special treatment that watershed solids are given in dredging simulations?
- Simulate MNA 2009 two ways for TIP
  - With baseline assumptions (50% erosion flux rule)
  - With special assumptions for dredging simulations (averaging normal and high adsorption runs)

➢ Special assumption increases export by ~13%

- This part of dredging model needs attention

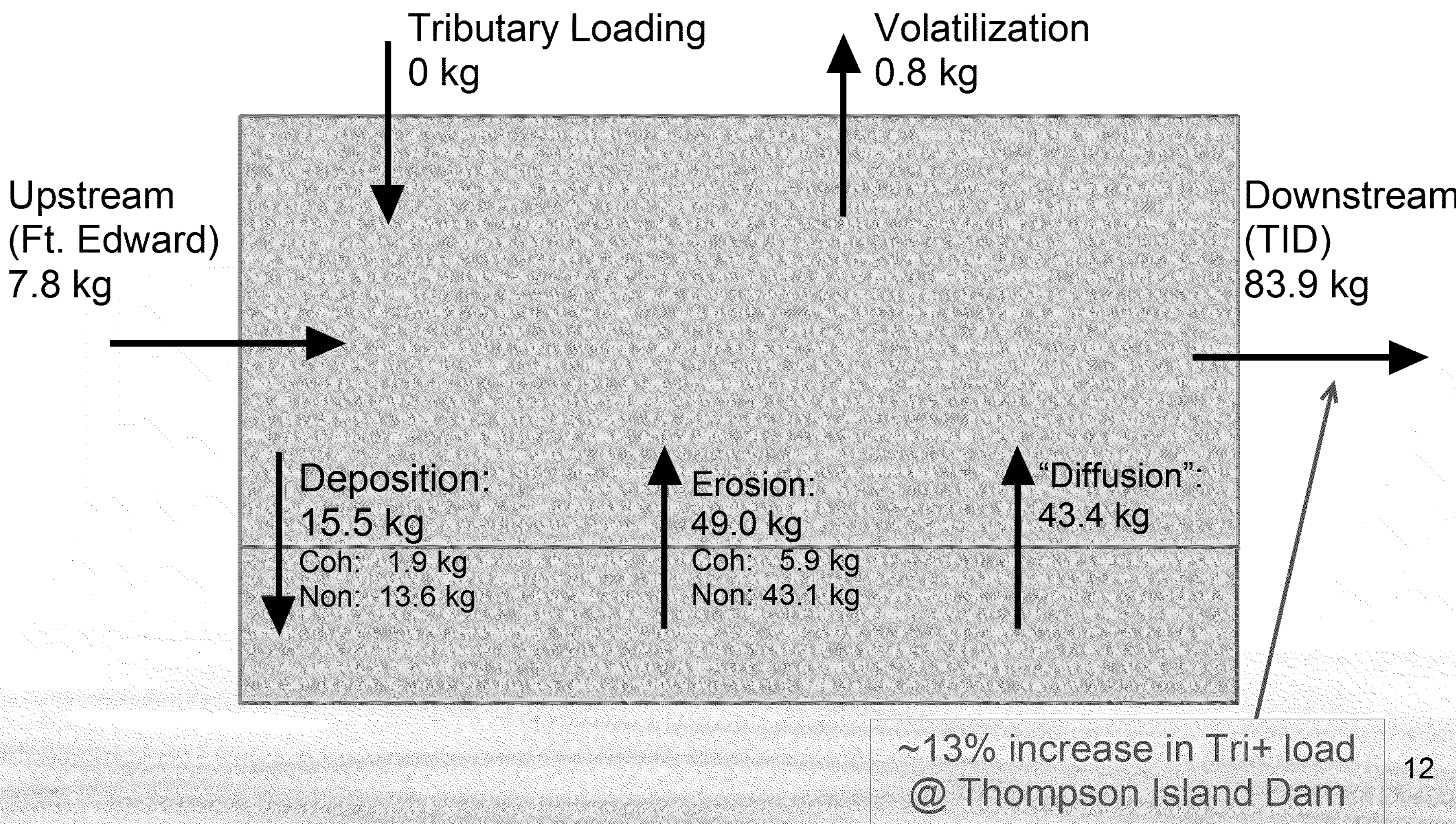
Mass Balance: MNA 2009 Tri+ (R8)

(with 50% rule for non-cohesive resusp.)



Mass Balance: MNA 2009 Tri+ (R8)

(normal/high Koc averaged)



Dredging Simulations – Assumptions Versus Nearfield Data

- Model assumptions ensure that water column PCB near dredging must partition as follows:
  - Particulate fraction > 0.5, dissolved fraction < 0.5
- Data from two Phase 1 nearfield studies show:
  - Transects: Particulate ~ 0.3, dissolved ~ 0.7
  - Special study: Particulate < 0.1, dissolved > 0.9
- Predicted fate of PCBs is expected to be very sensitive to this uncertain partitioning

Dredging Simulations – Spreadsheet Demonstration

- To demonstrate, assume:
  - Phase 1 export at Waterford = 200 kg total PCB
  - For volatilization:
    - Mass transfer coefficient = 0.3 m/d
    - Average water depth = 3 m
    - Time of travel to Waterford = 2.5 days
    - Additional loss at dams = 5% of resuspended PCB
  - No particulate PCB export at Waterford
- Allow particulate fraction to vary from 0.1 to 0.5

Dredging Simulation – Uncertain Fate of Resuspended PCBs

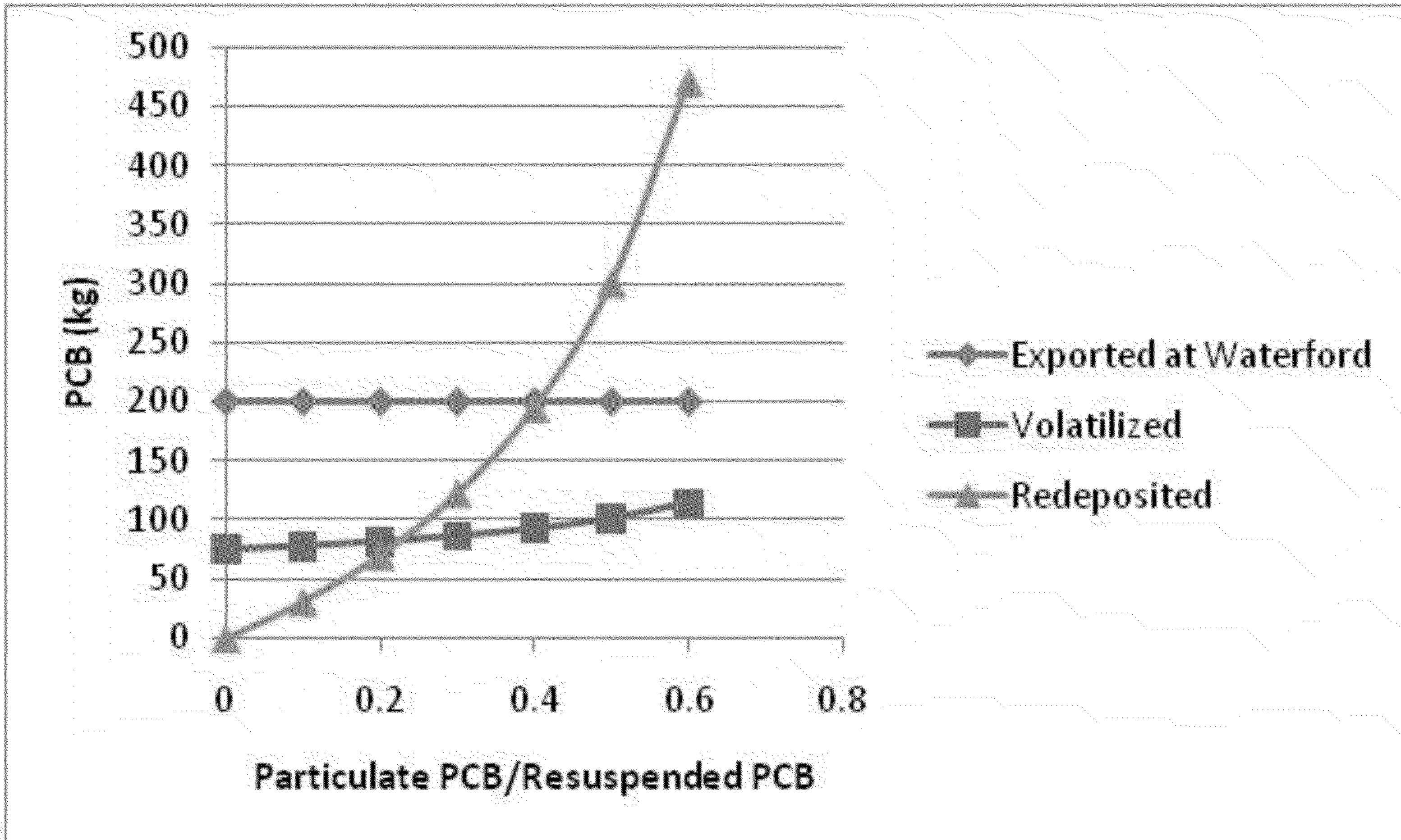
- Results under these assumptions:

~ 2/3 of dissolved fraction is exported (= 200 kg)

~ 1/3 volatilizes in Upper Hudson (~100 kg)

- Particulate fraction determines redeposition
  - If particulate fraction is 0.5, then about 300 kg must be in particulate form (& redeposit)
  - Lower particulate fractions in Phase 1 nearfield data imply lower resuspended PCB mass in particulate form, less PCB redeposition

Dredging Simulation – Importance of Uncertain Particulate Fraction



Dredging Simulations – Data Are Insufficient to Constrain Model

- Fate of resuspended PCBs and resulting recovery of Upper Hudson depends on:
  - The true nearfield dissolved/ resuspended split
  - How much resuspended PCB truly volatilizes
  - How much particulate PCB is exported to the Lower Hudson
- EPS Peer Review Panel: "there are insufficient data specific to near-field PCB releases to support appropriate calibration and validation of any model"

Dredging Simulations – What is Needed

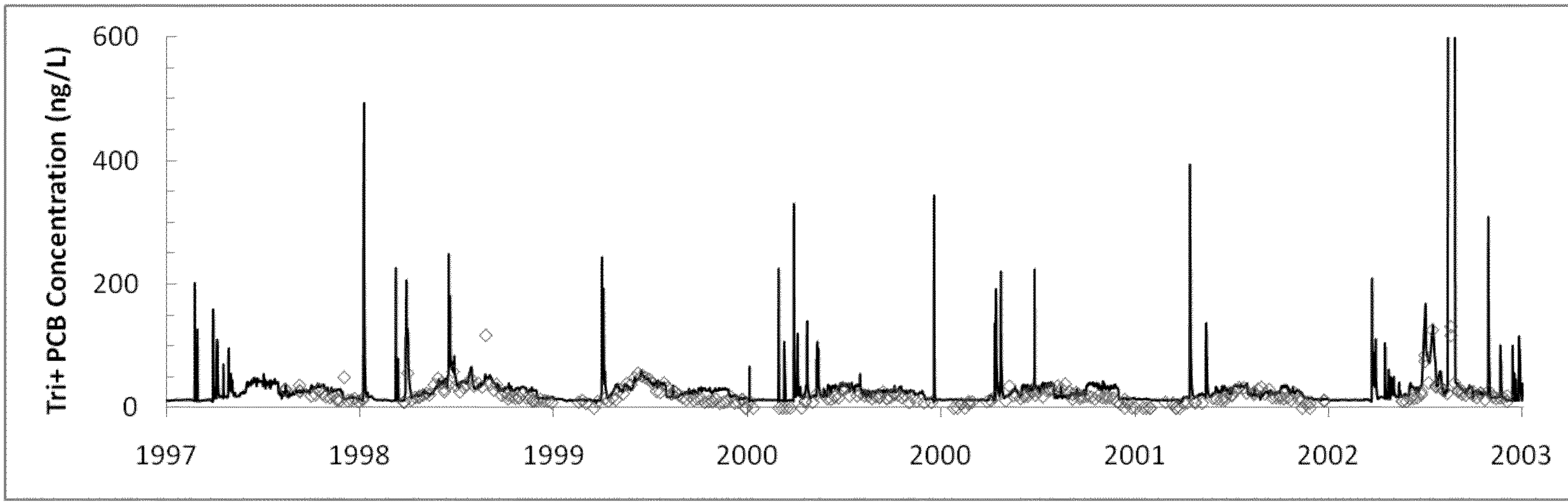
- Better data on the fate of resuspended PCBs from Phase 2/Year 1 data are needed before recovery can be forecast with confidence
- Mechanism may need reformulation in model
  - If there is resistant desorption, it should be assumed only for resuspended solids
  - May need to explicitly model resistant desorption, by representing resistant-phase PCBs as a separate state variable in the water column

Baseline Model Evaluation

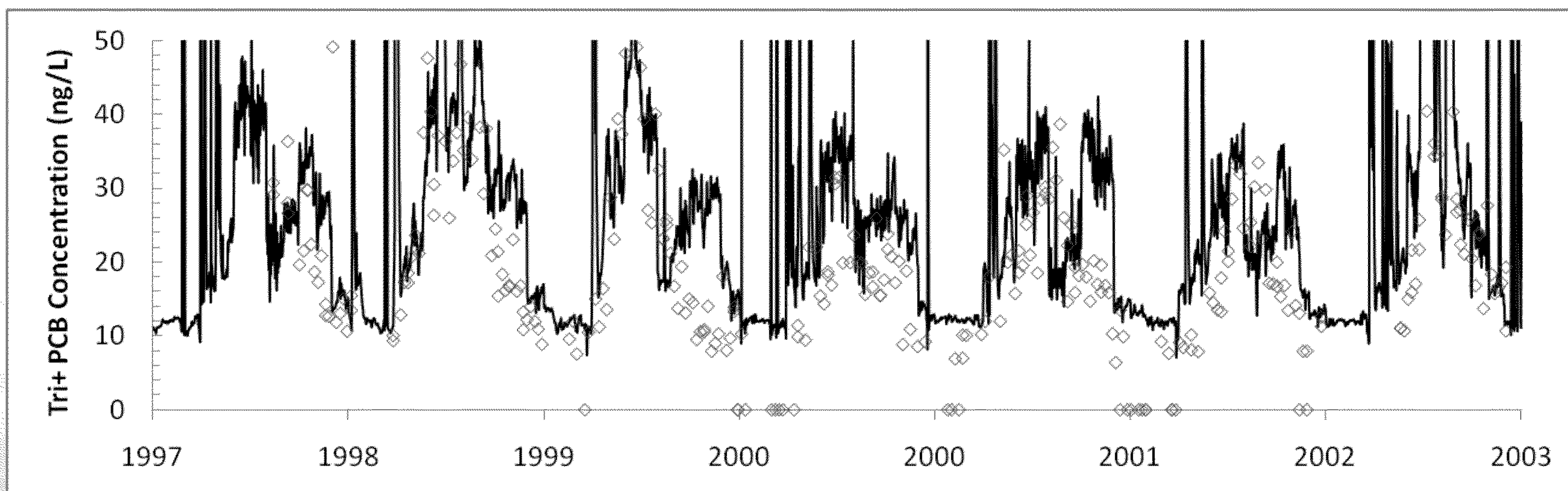
- More formal model-data comparisons are needed to ensure accuracy
  - Standard report presentation is time series plot
  - Visual presentation not sufficient to verify
    - Whether model has overall biases
    - Whether biases occur at high or low flow
    - Whether simulated long-term trends match data

Tri+ PCB at TID from 1997 to 2003

Figure 6-30



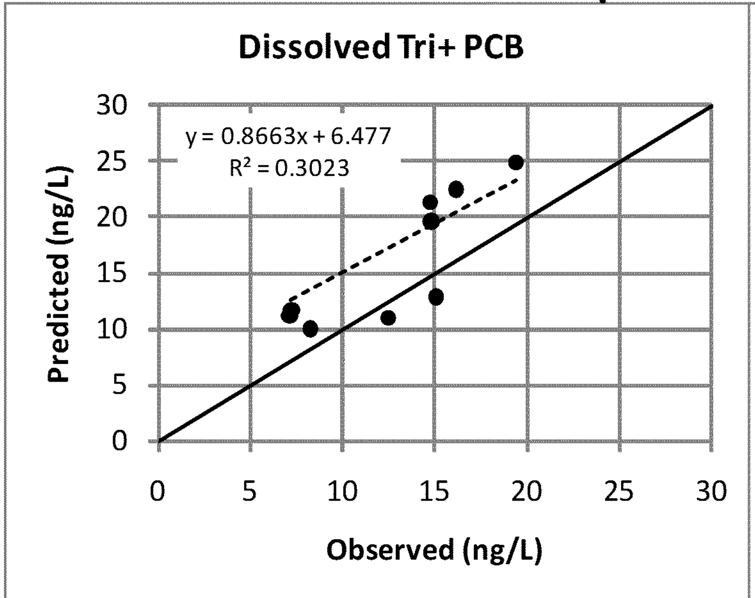
New Y-axis Range

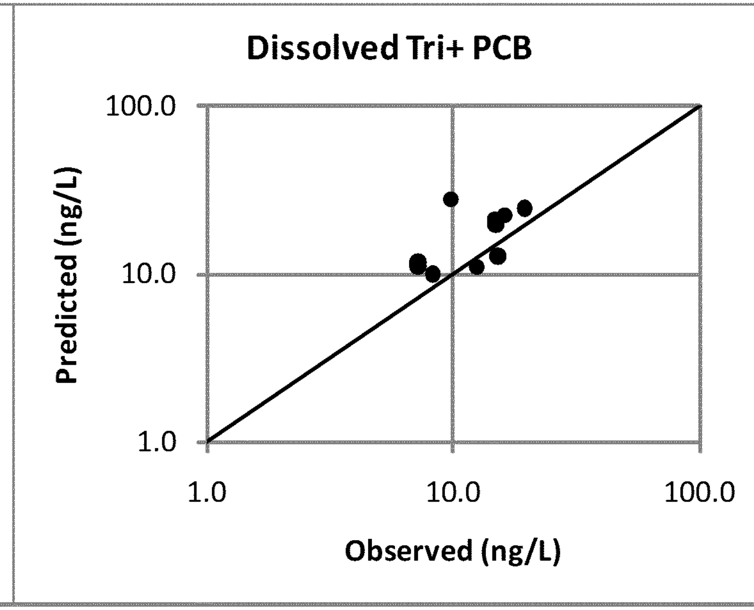


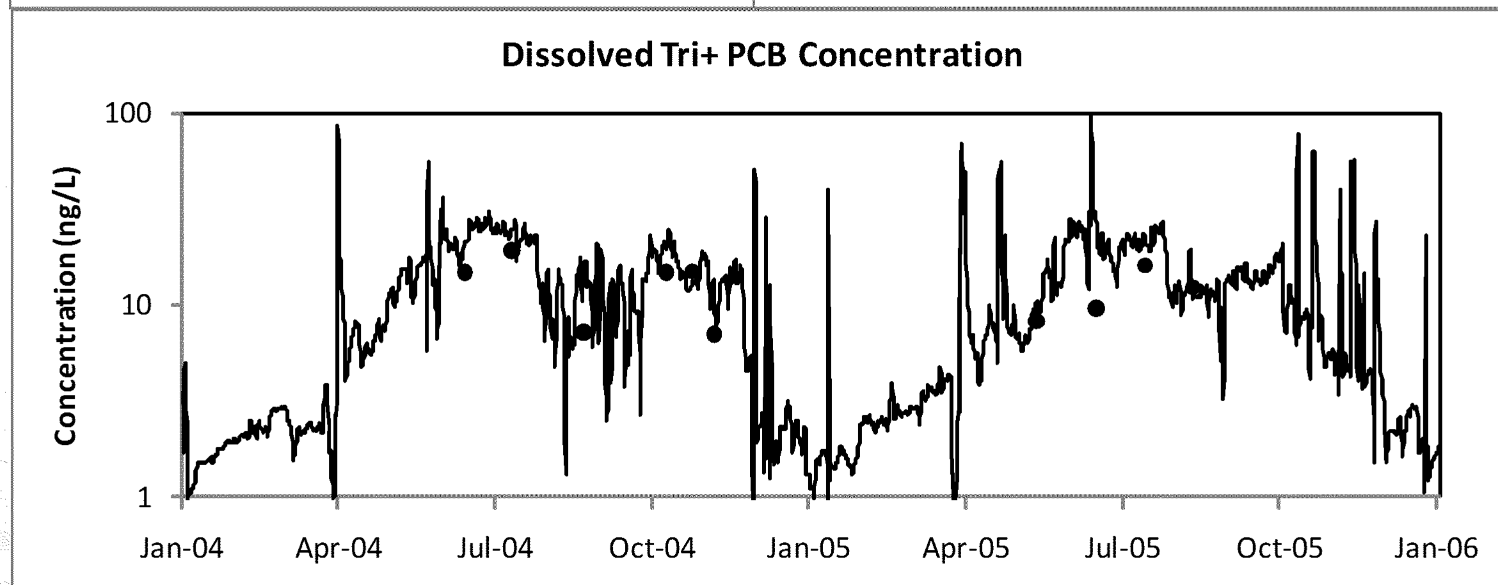
Baseline Model Comparisons

- Formal comparisons will employ:
  - Quantitative analysis of prediction error
  - Model-data scatter plots
  - Cumulative frequency distributions, model vs. data
  - Long-term time trend comparisons, model versus data
- Add dissolved/particulate PCB and TSS comparisons

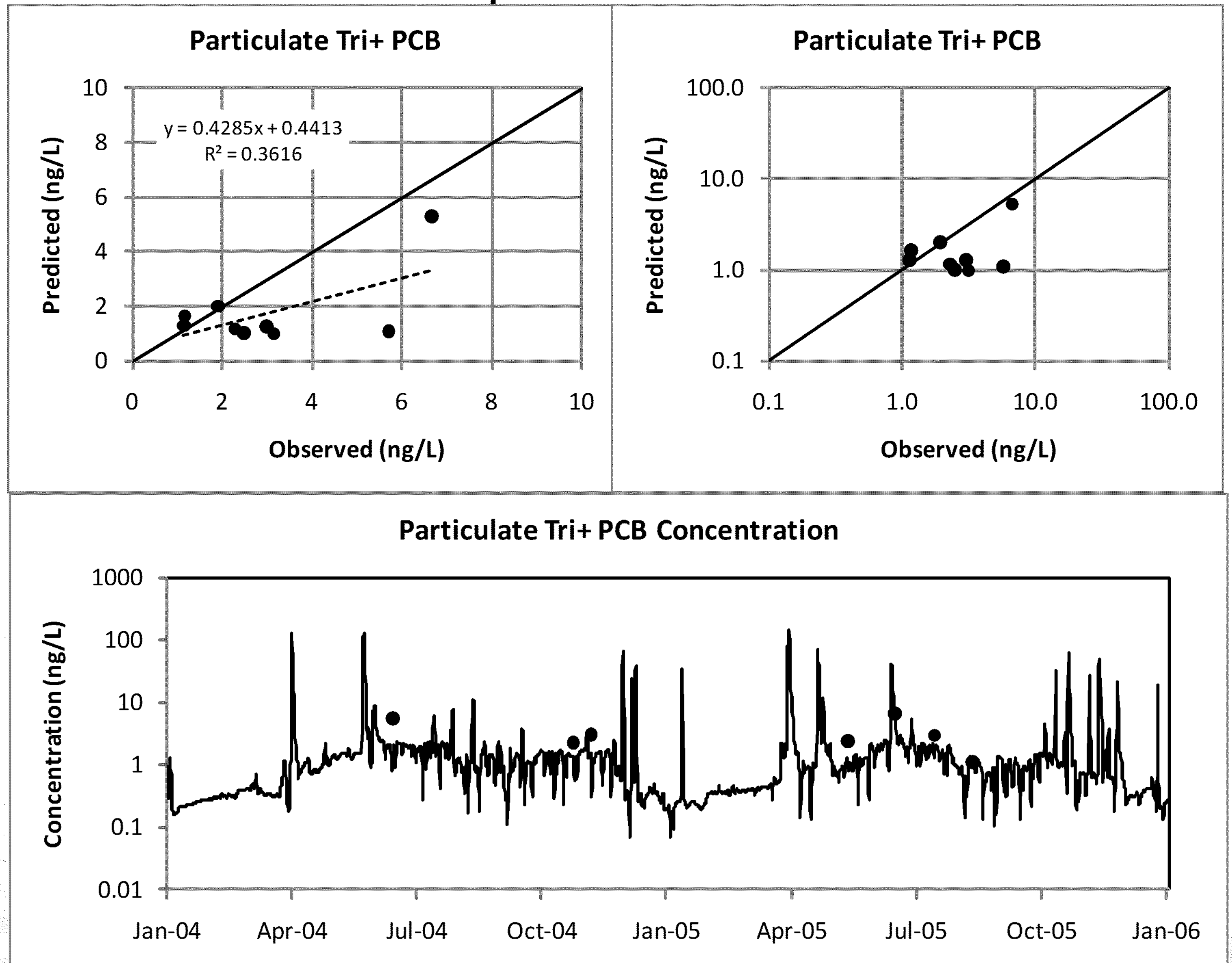
Thompson Island Dam



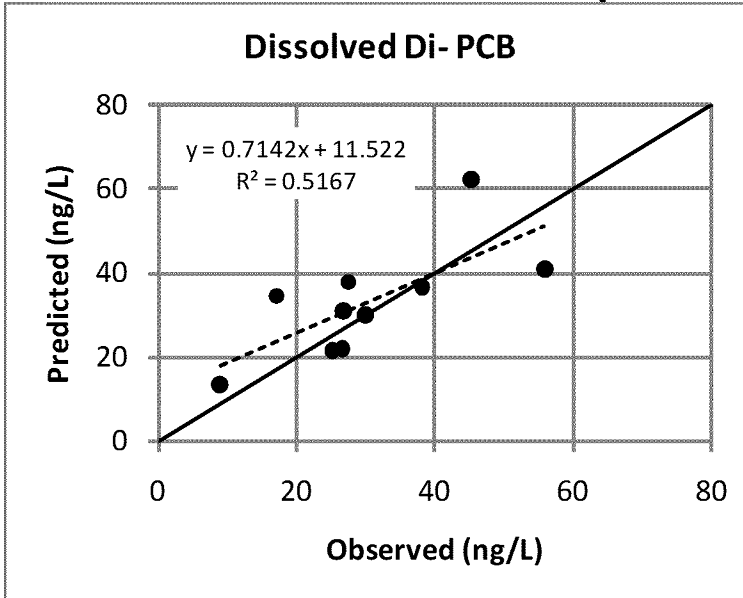


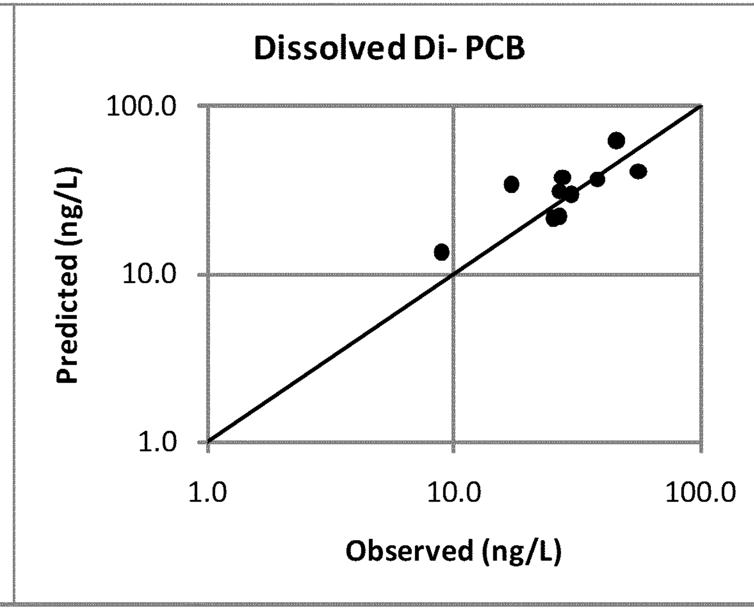


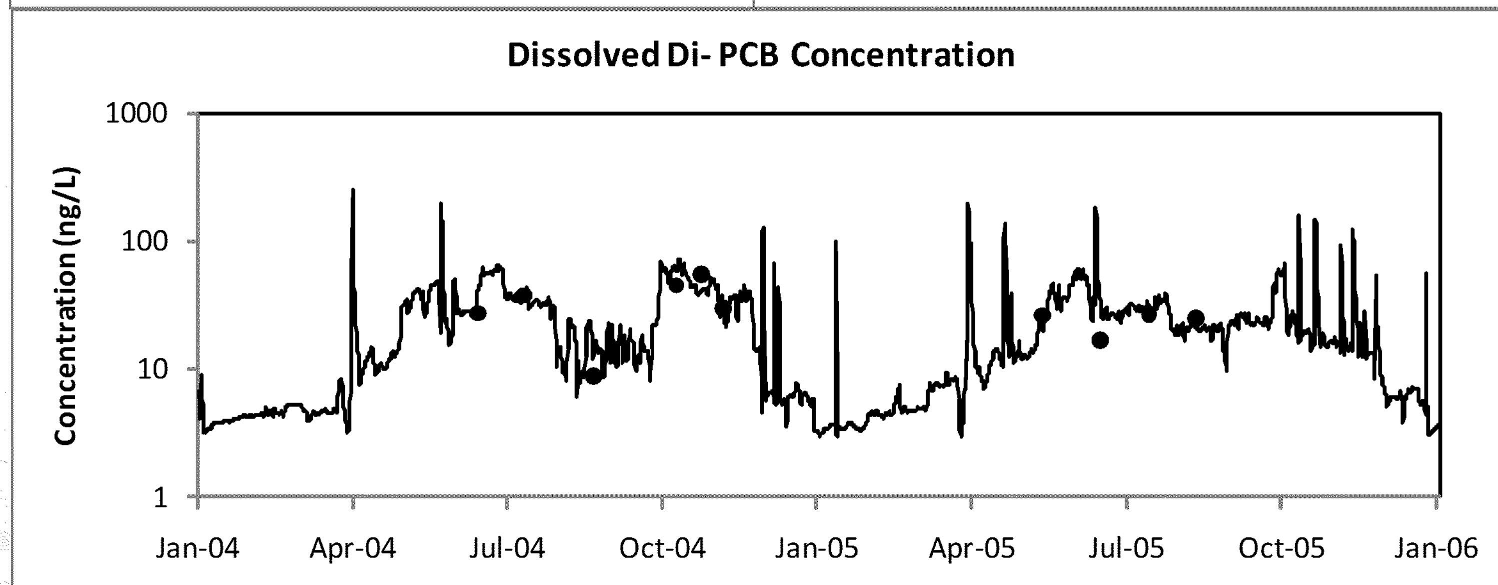
Thompson Island Dam



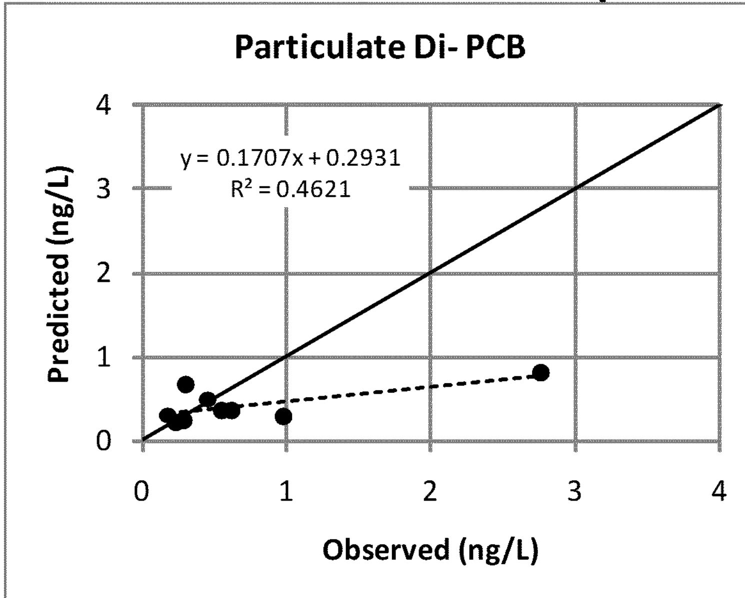
Thompson Island Dam

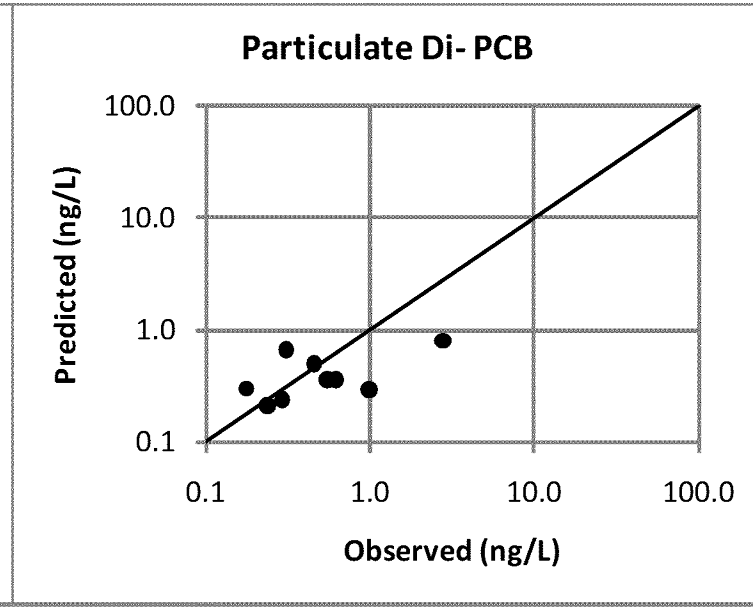


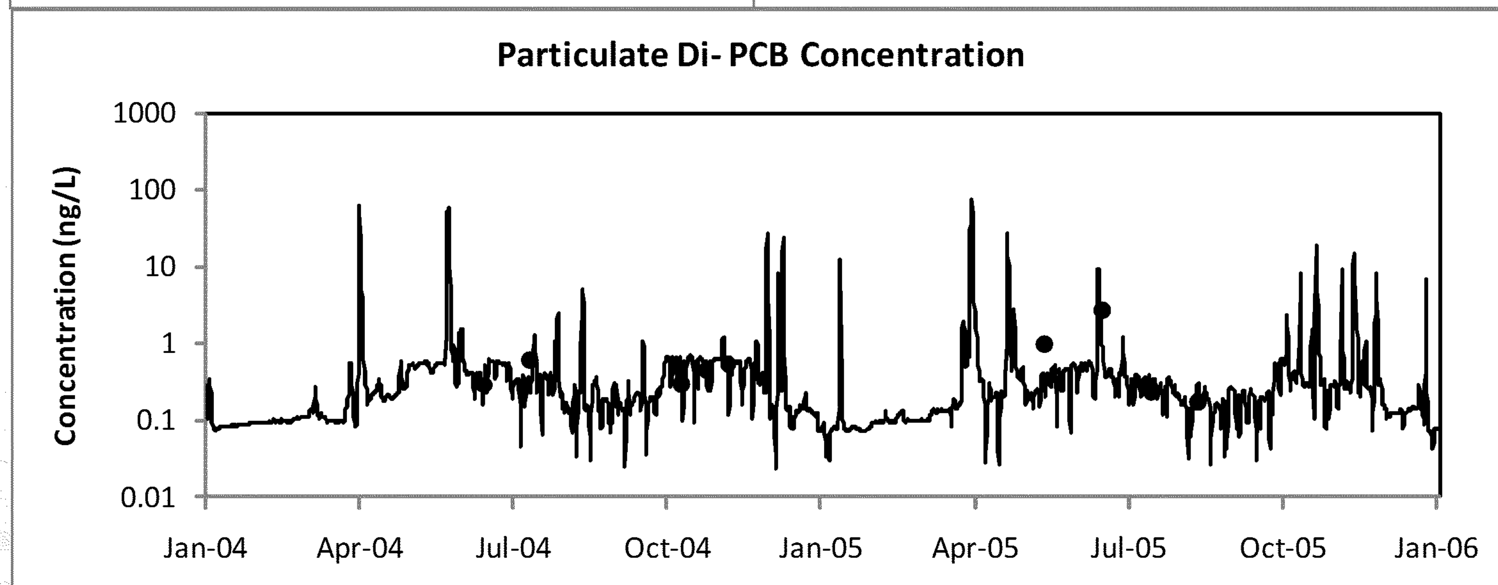




Thompson Island Dam







Uncertainty in Long Term Trends – PCB Fate Model

- Long-term fate model calibration is tied to 1977 - 2004 sediment trend
  - 1977 sediment concentrations highly uncertain
  - Only visual comparisons of predicted water column and fish tissue trends to data are shown
  - More rigorous model-data trend comparisons are needed, especially for rich water column dataset

Uncertainty in Long Term Trends – Sediment Transport Model

- Sediment transport model is critical to predicted recovery
  - Long-term net burial sequesters surface contamination
  - Mixing keeps PCBs near surface
- Predicted long-term burial rates are not compared to data (Cs-137 or bathymetry)
- PCB calibration feeds back to mixing but not to burial rates (settling - resuspension)

Sediment Transport – Model Versus Data

- Extensive grain size data are reported in Chapter 5 of June 2010 model report
- Effective diameters (µm), data versus model:

Class Limits	<62	62-250	250-2,000	> 2,000
Site Data	27	130-174	546-720	1,645-7,839
Model	30	90	1,500	8,000

- Classes 2 & 3 are calibrated outside of their data ranges

Significance of Model Data Difference in Grain Sizes

- Class 2 solids (fine sands) are more easily resuspended in model than if based on data
- Class 3 solids (med-coarse sands) are better able to armor the bed than if based on data
  - Is the model moving too much fine sand and too little coarse sand during simulated events?
  - If model had bed load would this be necessary?
- Diagnostic: Rerun 1994 flood event with data-based particle sizes

AQ R8 Sed Trans Calibration Plot from June 2010 Model Report

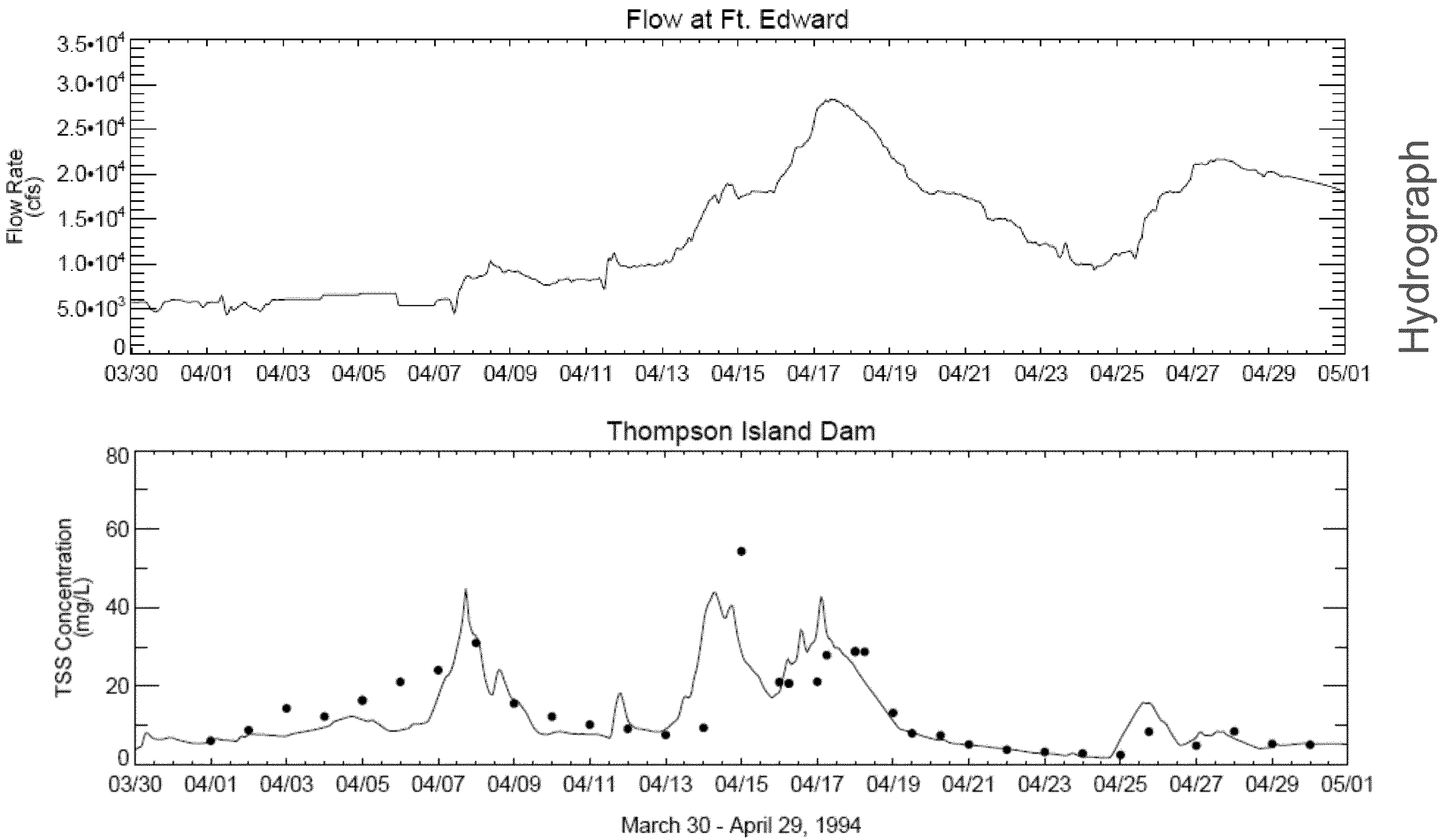
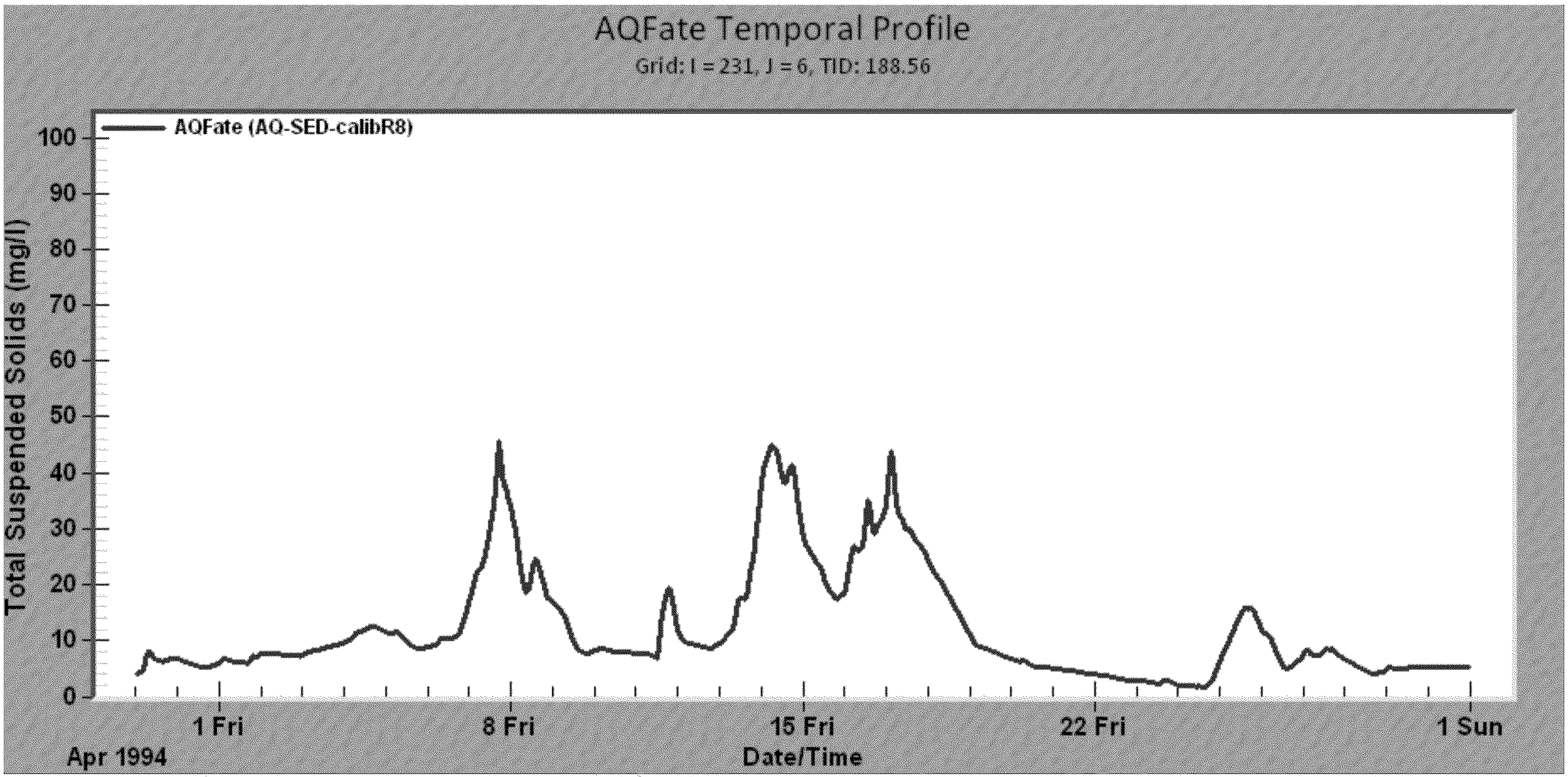


Figure 5-18

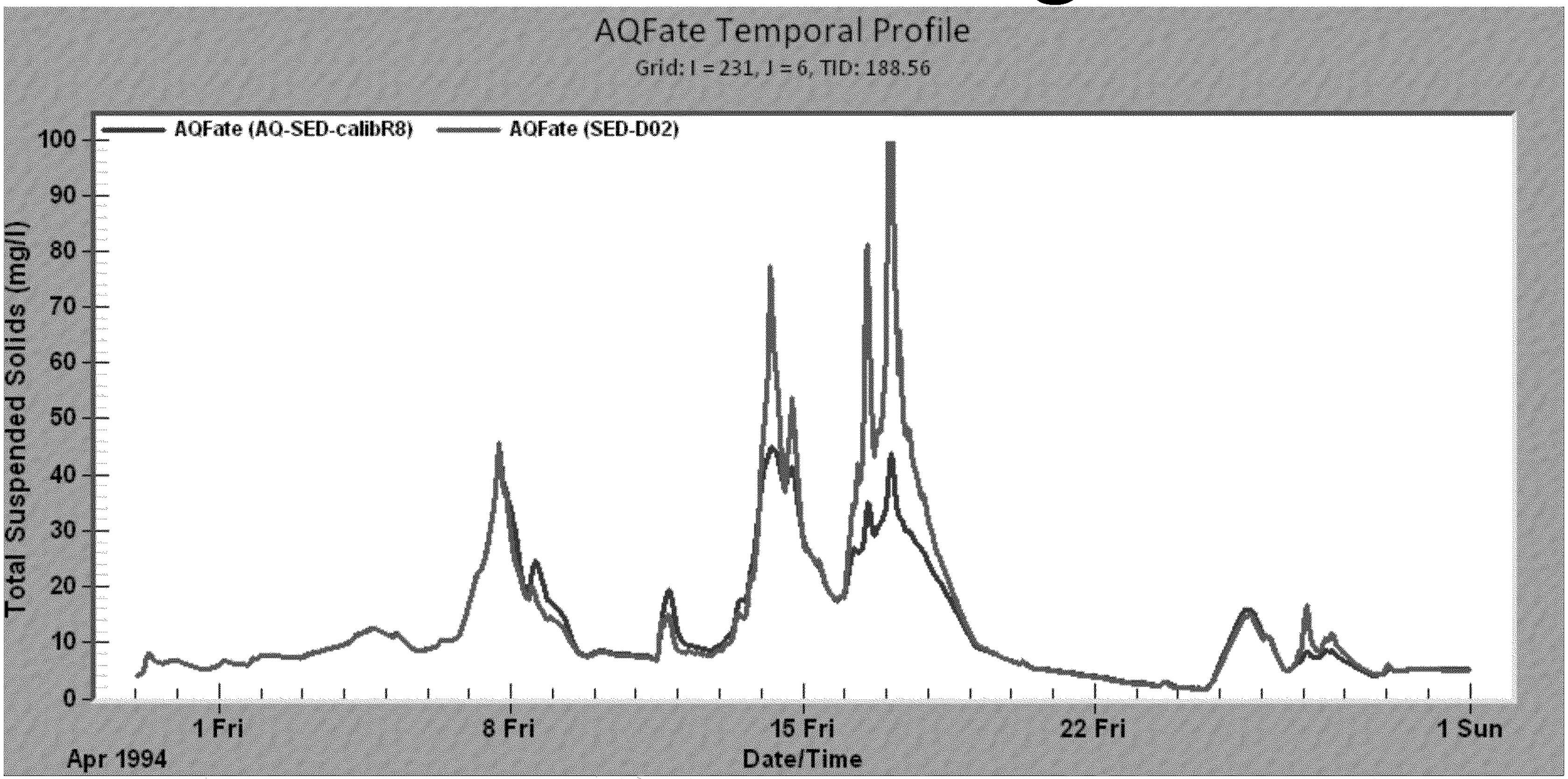
Comparison of Predicted (line) and Measured (symbols) Suspended Sediment Concentrations at Three Locations in the Thompson Island Pool During the 1994 Flood

Note: Top panel Shows Measured Flow at Fort Edward.

Benchmark Run Reproduced AQ Output



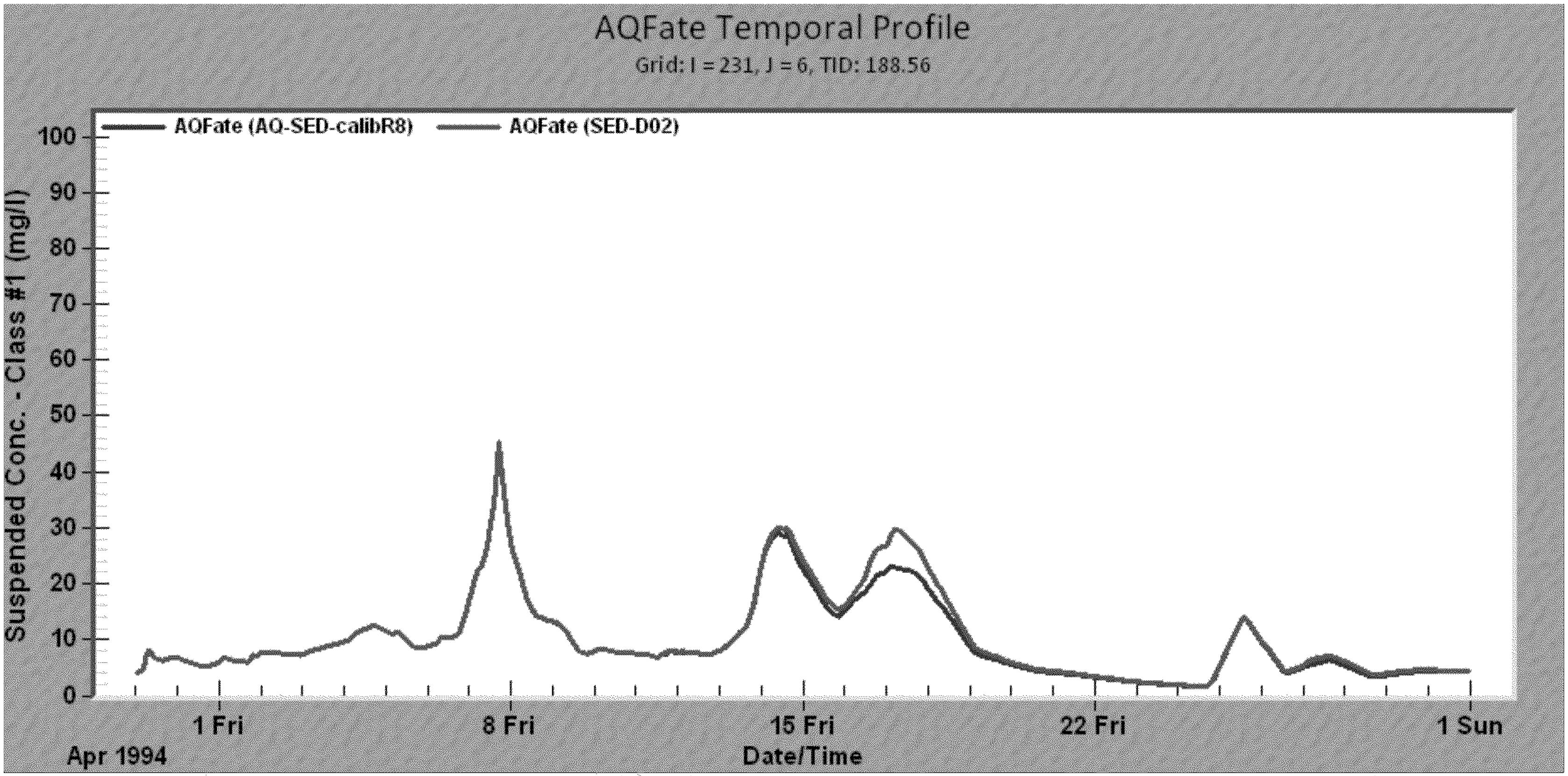
With Data-Based Inputs, Peak Event TSS is Higher



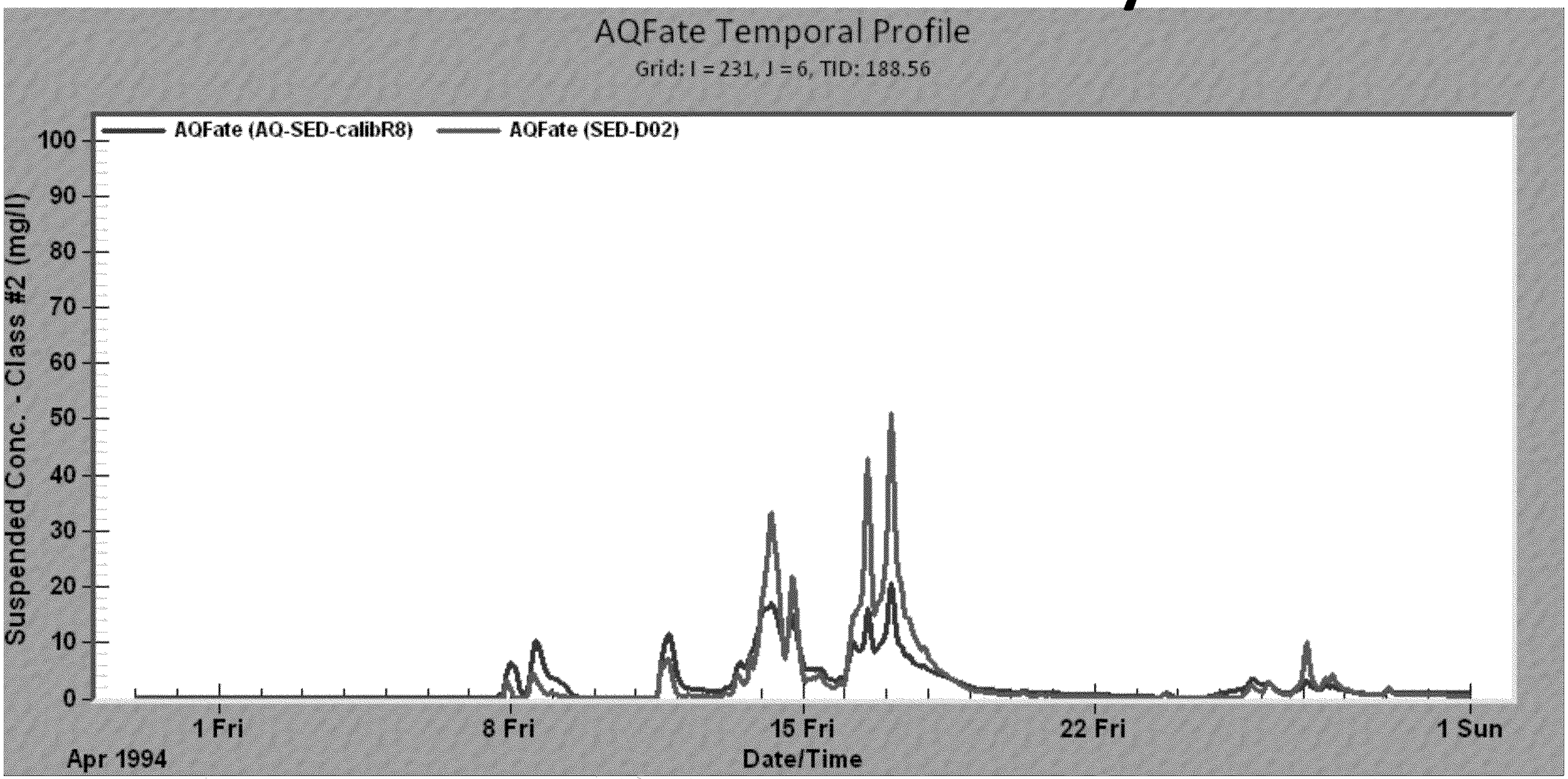
Reach 8 TSS at Peak Event Flow (April 17): AQ Calibrated Effective Particle Diameters (d2 = 90µm, d3 = 1500µm)

TSS at Peak Event Flow (April 17): Data-Based Effective Particle Diameters (d2 = 145μm, d3 = 670μm)

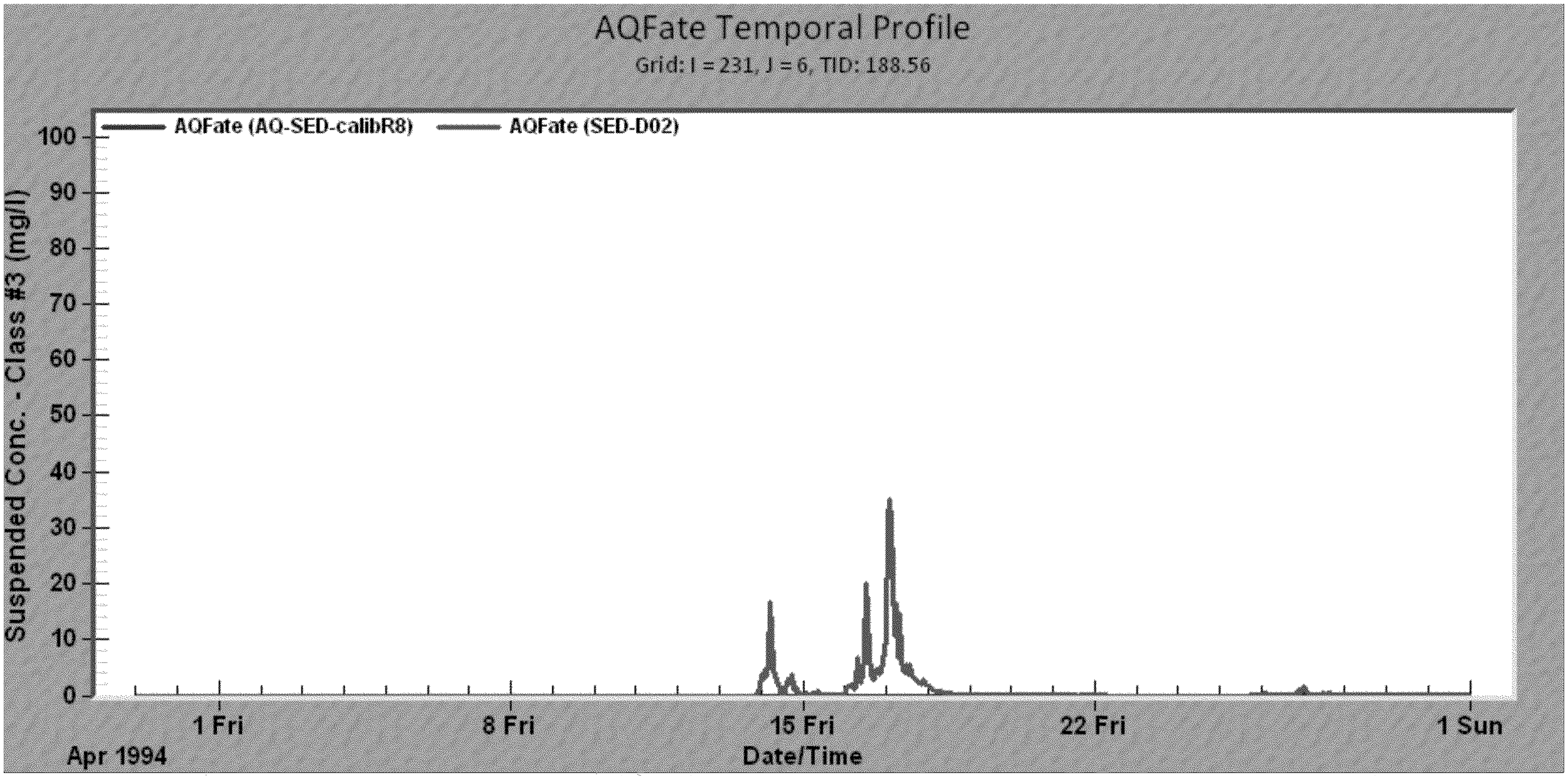
Suspended Silt/Clay: Differences are Small



Fine Sand: Larger Size Based on Data – Less Erodible Early in Event



Coarse Sand: Smaller Size Based on Data – Erodes at Peak Flow

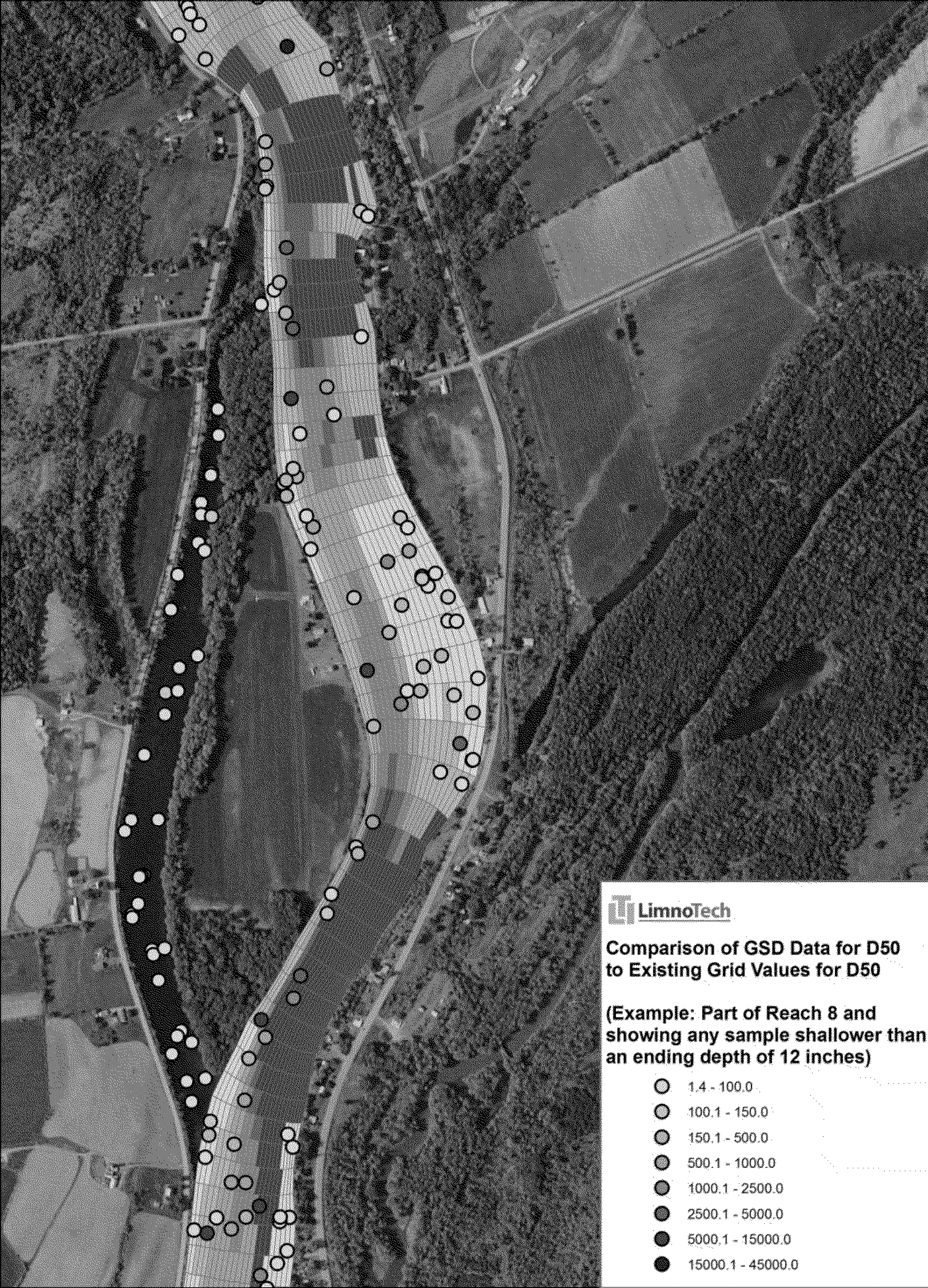


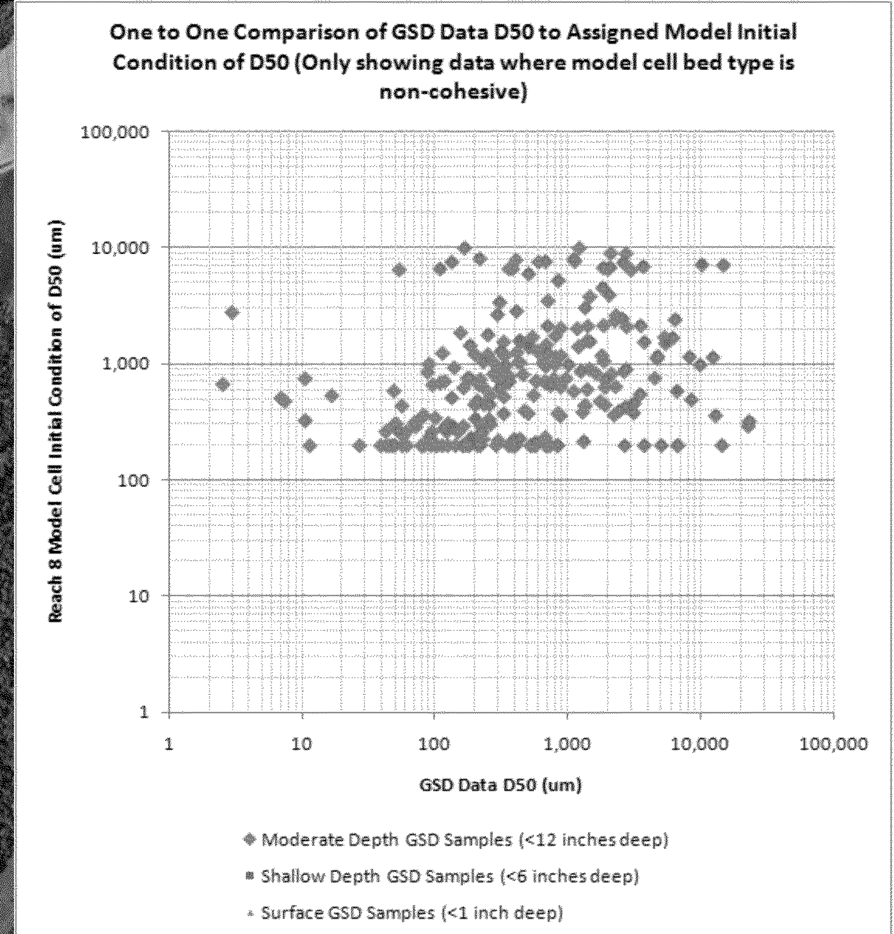
Significance for Model Evaluation

- Need to understand why model had to deviate from data to hit calibration targets
  - Are other processes omitted?
  - If they were included, how different would the simulations be?
  - Implications for uncertainty bounds of forecasts
- Currently conducting model-data comparisons on suspended solids to better evaluate sediment transport model accuracy

Input Check – Sediment Transport

- Median grain size d_{50} was estimated using shear stresses from hydrodynamic model
  - Model then uses it to set the fraction of solids that are small enough to be suspended
  - Chapter 5 of June 2010 report says that these were adjusted until predicted and measured distributions of d_{50} "were in general agreement"
- Checking a portion of Reach 8 shows weak model-data relationship over space


LimnoTech
Comparison of GSD Data for D50 to Existing Grid Values for D50
(Example: Part of Reach 8 and showing any sample shallower than an ending depth of 12 inches)
1.4 - 100.0
100.1 - 150.0
150.1 - 500.0
500.1 - 1000.0
1000.1 - 2500.0
2500.1 - 5000.0
5000.1 - 15000.0
15000.1 - 45000.0


One to One Comparison of GSD Data D50 to Assigned Model Initial Condition of D50 (Only showing data where model cell bed type is non-cohesive)
Reach 8 Model Cell Initial Condition of D50 (um)
100,000
10,000
1,000
100
10
1
1
10
100
1,000
10,000
100,000
GSD Data D50 (um)
Moderate Depth GSD Samples (<12 inches deep)
Shallow Depth GSD Samples (<6 inches deep)
Surface GSD Samples (<1 inch deep)

Path to Peer Review/Acceptance

- Baseline model
  - Complete rigorous comparisons of model to data
  - Conduct diagnostics to understand uncertainties
  - Strengthen model using existing data
- Dredging model
  - Improve representation of resuspension
  - Collect data in Phase 2/Year 1 for calibration
  - Demonstrate value to support Phase 2/Years 2+

Opportunities for/Value of Collaboration

- Faster route to peer review and acceptance
  - Have model ready for use in 2012
- Agreement on terms and objectives of collaboration would help to focus effort
- Consistent with recommendations of EPS panel
  - Panel offered to provide continuing oversight after adding a modeler

Summary

- Systematic evaluation of full suite of models is underway
- Acceptance of models for management will require rigorous comparison to data and modifications as needed
  - Baseline model to existing data
  - Dredging model to Phase2/Year 1 data
- Collaboration can greatly speed that process